ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/291331547

Computational model for Bangla text to Braille translation

Conference Paper · February 2011

CITATIONS
2

READS
626

5 authors, including:

Fakhruddin Muhammad Mahbub ul Islam
St. Francis Xavier University
7 PUBLICATIONS **58** CITATIONS
SEE PROFILE

Samiul Azam
United International University
16 PUBLICATIONS **128** CITATIONS
SEE PROFILE

Md Osman Gani
Marquette University
12 PUBLICATIONS **60** CITATIONS
SEE PROFILE

Ahamad Imtiaz Khan
Constellation Software
18 PUBLICATIONS **34** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project An offline and online-based Android application “TravelHelp” to assist the travelers visually and verbally for Outing View project

Project Satisfaction minining for students in university View project

All content following this page was uploaded by Ahamad Imtiaz Khan on 19 January 2017.

The user has requested enhancement of the downloaded file.

Computational Model for Bangla Text to Braille Translation

Fakhruddin Muhammad Mahbub-ul-Islam, Ahamad Imtiaz Khan, Md Osman Gani, Samiul Azam
Department of Computer Science and Engineering
Military Institute of Science and Technology
Dhaka, Bangladesh
mahbub.mist@yahoo.com, aumi_aik@yahoo.com, oscse23@gmail.com, sami.mist@yahoo.com

Syed Akhter Hossain
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
aktarhossain@yahoo.com

Abstract—**The Braille is a system that is widely used by blind people to read and write. Visual impaired persons in Bangladesh are deprived of education and modern technologies due to the lacking of textbooks and other reading materials written in Bangla Braille. This is indirectly widening the knowledge gap and disparity within the society. In order to bring a radical change into this scenario, an automated system for facilitating machine translation of Bangla text to readable and recognizable Braille symbol is needed. In this paper, a detail research on Bangla Braille has been accomplished and necessary rules as well as grammatical conventions are analyzed. Based on the research and analysis, this study illustrates a computational model using Deterministic Finite Automata (DFA), which can be extended to an automated system for Bangla text to Braille translation. The designed DFA demonstrated acceptable conversion.**

Keywords-braille cell; bangla braille; conversion rules; computational mode; deterministic finite automata; view model

I. INTRODUCTION

Bangla is the national language of Bangladesh while it is also the state language of the Indian state of West Bengal. It is one of the most spoken languages (ranking sixth) in the world, with nearly 300 million total speakers [1]. According to [2] about 0.75 million people are blind in Bangladesh. Blind person cannot read text written on a plain paper. They use Braille system in their writing and reading. Braille is a universal code for mapping character sets of various languages to Braille cells. Each Braille cell consists of 6 dots. Series of Braille cells are embossed on paper so that visually impaired persons can read them by touching the raised dots of the cells using their fingers. Like other languages Bangla has own representation in Braille system.

There are few schools, institutes and organizations for blind people in Bangladesh where they can get their education. But they always face one problem that is lack of necessary reading materials, books and notes in Braille written format. Rewriting of books into Braille language is done by certified Braille experts. But this process takes lots of time and cost. For this reason books written in Braille language are insufficient. We consider this situation as a major problem in the blind's education. This problem can be solved by an automated translation system which can convert electronic Bangla books to Bangla Braille books. For developing this system a model is proposed in this paper.

This paper is organized as follows: Section II explains related works about text to Braille conversion of different languages other than Bangla. Introduction of Braille and Bangla is described in section III. In section IV a computational model is proposed. In section V, a view model is shown which illustrates the proposed computational model. Finally the conclusion and discussion is outlined in Section VI.

II. LITERATURE REVIEW

Liblouis is an open source text to Braille and Braille Back translator [3]. It supports contracted and uncontracted Braille for many languages such as: Arabic, Armenian, Bulgarian, Chinese, Croatian, Czech, Danish, Dutch, English, Esperanto, Estonian, Finish, French, Gaelic, German, Greek, Icelandic, Italian, Lithuanian, Norwegian, Polish, Portuguese, Romanian, Russian, Slovakian, Spanish, Swedish, Turkish, Vietnamese and Welsh. New languages can be easily added through Tables that support a rule or dictionary based approach. It has also tools for testing and debugging tables.

"BrailleTrans for Java" is an open source solution for text to Braille translation which is implemented in Java [4]. It is built on an existing system which was implemented on C and includes a finite state machine with left and right context matching and a set of translation rules. It can translate different languages to different grades of Braille contraction. It also allows both forward and backward translation i.e. text to Braille and Braille to text translation.

Hardware-based Text to Braille translation is a system which performs faster and is able to generate greater throughput than Software-based Text to Braille translation system [5]. This system is based on Field Programmable Gate Arrays (FPGAs) for fast text to Braille translation.

An automated system exists for English and Hindi text to Braille representation which is implemented in Java [6]. It uses a chart as database. This database is used for mapping corresponding Braille representation. Then it implements the mapping for translation. It can translate both English and Hindi to corresponding Braille.

Quranic Braille system translates Quranic verse to Braille symbol. It is implemented in Visual Basic [7]. It can translate the special vibration for the Quran. But it is limited for the (Noun + Scoon) vibrations.

TextBraille is a user-friendly Braille Transcription software for Indian languages for Windows. The software transcribes the text files to the Braille format, shows the text file as well as the Braille file side by side. It can translate various Indian languages to Braille such as Hindi, Bengali, Assamese, Oriya, Marathi, Gujarati, Tamil, Telegu, Kannada, Malayalam, Punjabi [8]. It can translate Bangla to Braille but this software does not fulfill our purpose because Bangla to Braille conversion rules for Indian Bangla and Bangladeshi Bangla is different. There are also others research works related to text to Braille conversion in different languages [9] [10] [11][12].

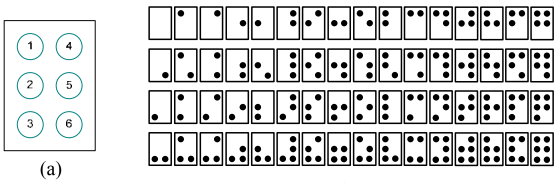
III. BANGLA AND BRAILLE

A. *Bangla*

In Bangla alphabet there are 11 independent vowels, 10 amongst them can also be used as dependent vowels and 39 consonants. Independent vowels are অ, আ, ই, ঈ, উ, ঊ, ঋ, এ, ঐ, ও, ঔ; Dependent vowels are া, ি, ী, ু, ূ, ৃ, ে, ৈ, ো, ৌ and Consonants are ক, খ, গ, ঘ, ঙ, চ, ছ, জ, ঝ, ঞ, ট, ঠ, ড, ঢ, ণ, ত, থ, দ, ধ, ন, প, ফ, ব, ভ, ম, য, র, ল, শ, ষ, স, হ, ড়, ঢ়, য়, ৎ, ং, ঃ, ঁ.

B. *Braille*

The Braille system is a universal way for the blind people to read and write. A blind Frenchman Louis Braille invented this system in 1821. The unit of Braille is a Braille cell. Each Braille cell consists of 6 dot positions, put together in a rectangle. Six dot positions are assembled in that rectangle in two columns, three dots in each. Dot positions 1-2-3 are in first column and 4-5-6 are in second column which is shown in Figure 1(a). A dot position can have two states. It may be raised or not raised. So depending on these two states of six dot positions, 64 permutations are possible [13]. Figure 1(b) shows all 64 Braille cells. An example can be given by naming the raised dot positions. For example if it is said like dots 1-2-3-5, it describes dots 1, 2 and 3 are raised in first column and dot 5 is raised in second column. This particular permutation indicates English letter *R*.



(a)

(b)

Figure 1: (a) A Braille cell, (b) 64 Braille cells

C. *Mapping of Bangla charater to Braille*

Braille is not a language. It is a code for mapping character sets of various languages to its 64 fixed permutations. Mapping of character sets of a language to Braille symbols is called character mapping. Different countries developed their own standard character mapping for their language. This character mapping is not necessarily one-to-one because some languages may have more than 64 letters in their alphabet. So possible character mapping can be one-to-one and many-to-one. For example

Bangla letter ‘ক’ → ⠅ (One-to-one)
Bangla letter ‘হ’ & digit ‘৮’ → ⠓ (Many-to-one)

Each letter of Bangla alphabet, numeric and punctuation has corresponding Braille representation. Bangla letters and their corresponding Braille representations are given in Table I, II and III.

TABLE I. ONE TO ONE CHARACTER MAPPING

Bangla	**Braille**	**Bangla**	**Braille**	**Bangla**	**Braille**
ক	⠅	ধ	⠮	ড়	⠻
খ	⠭	ন	⠝	ঢ়	⠷
ঘ	⠣	প	⠏	য়	⠢
ঙ	⠬	ম	⠍	ৎ	⠂⠞
ছ	⠡	য	⠽	ঁ	⠄
ঝ	⠵	র	⠗	ং	⠰
ট	⠾	ল	⠇	ঃ	⠠
ঠ	⠺	শ	⠩	ক্ষ	⠟
ড	⠫	ষ	⠯	জ্ঞ	⠱
ঢ	⠿	স	⠎	;	⠆
ত	⠞	।	⠲	!	⠖
থ	⠹	‘	⠠⠦	=	⠰⠶
ভ	⠧	’	⠴⠄	*	⠔⠔
-	⠤	[	⠠⠶	]	⠶⠄

TABLE II. TWO TO ONE CHARACTER MAPPING

Bangla		**Braille**	**Bangla**		**Braille**
অ	১	⠁	চ	৩	⠉
আ	া	⠜	জ	০	⠚
ঈ	ী	⠔	ঞ	:	⠒
উ	ু	⠥	ণ	Number prefix	⠼
ঊ	ূ	⠳	দ	৪	⠙
ঋ	ৃ	⠐⠗	ফ	৬	⠋
ও	ো	⠕	ব	২	⠃
ঔ	ৌ	⠪	হ	৮	⠓
গ	৭	⠛	্	”	⠴
(	)	⠶	?	“	⠦

TABLE III. THREE TO ONE CHARACTER MAPPING

Bangla			Braille
ই	ি	৯	⠊
এ	ে	৫	⠑
ঐ	ৈ	/	⠌
,	.	’(lop)	⠂

D. *Grammertical conversion rules*

In conversion of Bangla text to Braille, one needs to deal with conjunctions, consonants, dependent and independent vowels, punctuations and numbers. Here all grammatical conversion rules are discussed with examples.

1) Replacing rule: In replacing rule each Bangla character is replaced by its corresponding Braille cell(s) [14]. Some uses of consonants, vowels (dependent and independent) and punctuations are given in Table IV.

TABLE IV. USES OF CONSONANTS, VOWELS AND PUNCTUATIONS

Bangla Word	Distribution	Braille Representation
বল	ব ল	⠃ ⠇
উৎস	উ ৎ স	⠥ ⠂ ⠞ ⠎
কাঠ	ক া ঠ	⠅ ⠜ ⠺
দৃঢ়	দ ৃ ঢ়	⠙ ⠐ ⠗ ⠷
কমা	,	⠂
সেমি কোলন	;	⠆

2) Inserting rule: In inserting rule a Braille cell is inserted as prefix. Other characters are replaced according to replace rule [14].

a) If there is “ই”, “উ” or “ও” after consonant and if “অ” is pronounced there then Bangla “অ” equivalent Braille cell dot 1 is inserted after consonant in Braille [14]. Examples are shown in Table V.

TABLE V. INSERTION OF “অ”

Bangla Word	Distribution	Braille Representation
বই	ব ই	⠃⠁⠒
রওনা	র ও না	⠗⠁⠕⠝⠜

b) Braille cell dot 4 is inserted before conjunctions having combination of two letters [14]. Examples are shown in Table VI.

TABLE VI. CONJUNCTION OF 2 LETTERS

Bangla Word	Conjunct	Distribution	Braille Representation
গ্রাম	গ্র	গ ্ র	⠈ ⠛ ⠗ ⠜ ⠍
পূর্ব	র্ব	র ্ ব	⠏ ⠳ ⠈ ⠗ ⠃

c) Braille cell dot 4, 6 is inserted before conjunctions having combination of three letters or four letters [14]. Examples are shown in Table VII.

TABLE VII. CONJUNCTION OF 3AND 4 LETTERS

Bangla Word	Conjunct	Distribution	Braille Representation
রাষ্ট্র	ষ্ট্র	ষ ্ ট ্ র	⠗⠜⠨⠯⠾⠗
স্বাতন্ত্র্য	ন্ত্র্য	ন ্ ত ্ র ্ য	⠈⠎⠃⠜⠞⠨⠝⠞⠗⠽

d) Two Bangla conjunctions have direct representation in Bangla Braille. So there is no need to use Braille cell dot 4 before them [14]. They are given in Table VIII.

TABLE VIII. TWO CONJUNCTIONS WITHOUT PREFIX

Bangla Word	Conjunct	Distribution	Braille Representation
কক্ষ	ক্ষ	ক ্ ষ	⠅ ⠟
জ্ঞান	জ্ঞ	জ ্ ঞ	⠱ ⠜ ⠝

e) Braille cell dot 3,4,5,6 is inserted before first digit of a number [14]. Examples are shown in Table IX.

TABLE IX. NUMBER RULE

Bangla number	Braille Representation
১২৩৪	⠼ ⠁ ⠃ ⠉ ⠙
১২.৩৪	⠼ ⠁ ⠃ ⠂ ⠉ ⠙

IV. PROPOSED COMPUTATIONAL MODEL

Based on study of Bangla language, Braille system and conversion rules between Bangla and Braille we propose a computational model to translate Bangla text to Braille. Our proposed computational model consists of four modules. Each module contributes to the whole translation process. The computational model is shown in Figure 2.

A. *Bangla Text*

Bangla text consists of Bangla alphabets, digits, punctuations, spaces and formatting characters. Later part of the computational model is applied on this Bangla text in iterative manner. Input Bangla text will be in Unicode and font independent.

B. *Rule Engine*

It is one of the major parts of computational model. The work of Rule engine is to generate token and corresponding rule number based on input Bangla text. A Deterministic Finite Automata (DFA) is proposed for developing the rule engine.

A DFA has a finite set of states and a finite set of input symbols. One state is designated the start state, and zero or more states are accepting states. A transition function determines how the state changes each time an input symbol is

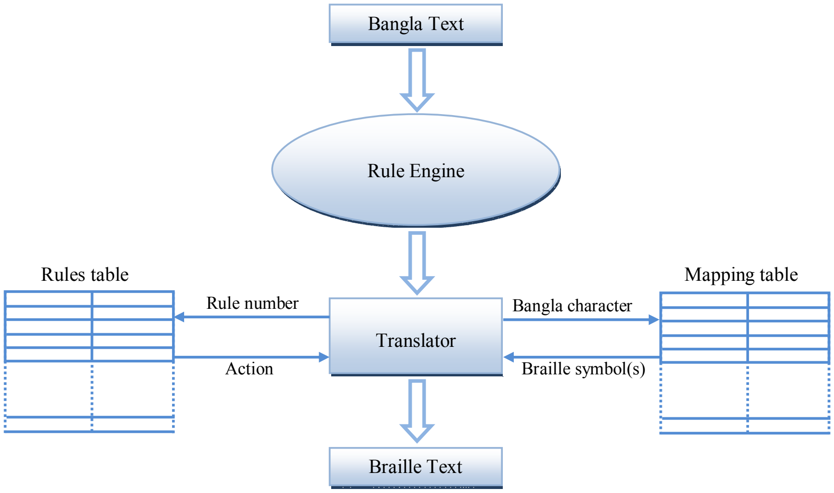


Figure 2: Computational Model for Bangla Text to Braille Translation

processed. The term ‘deterministic’ refers to the fact that on each input there is one and only one state to which the automaton can transition from its current state.

It is convenient to represent automata by a graph in which the nodes are the states and arcs are labeled by input symbols indicating the transitions of that automaton. The start state is designated by an arrow and the accepting states by double circles [15].

Proposed DFA is constructed according to Bangla text to Braille conversion grammar for tokenizing input Bangla text. Each accepting states indicates that a token has been found with applicable rules. The constructed DFA in “five tuple” notation is DFA, $A = (Q, \sum, \delta, q_0, F)$.

- Here A is the name of the DFA.
- Q is the set of states. It contains 26 states. $Q = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25\}$
- $\sum$ is input symbols. We have seven input symbols which are shown in Table X. Each of them represents set of Bangla characters.

TABLE X. INPUT SYMBOLS AND CHARACTERS

Input Symbols	Description	Bangla Characters
Vol_spe	Set of three independent vowels ই,উ,ও	ই, উ, ও
Vol	Set of independent vowels except ই,উ,ও and all dependent vowels	অ, আ, ঈ, ঊ, ঋ, এ, ঐ, ঔ, া, ি, ী, ু ূ ৃ, ে, ৈ, ো, ৌ
Con	Set of all Bangla consonants	ক, খ, গ, ঘ, ঙ, চ, ছ, জ, ঝ, ঞ, ট, ঠ, ড, ঢ, ণ, ত, থ, দ, ধ, ন, প, ফ, ব, ভ, ম, য, র, ল, শ, ষ, স, হ, ড়, ঢ়, য়, ৎ, ং, ঃ, ঁ
Num	Set of all Bangla digits	০, ১, ২, ৩, ৪, ৫, ৬, ৭, ৮, ৯
Pun	Set of all punctuation signs	, । ? ; : ‘ ’ “ ” - ! ( ) [ ]
Hos	Represents Bangla hasanta symbol	্
Dot	Represents decimal sign (.)	.

- q_0 is the start state of the DFA.

  q_0 = state 0.
- F is the set of accepting states. It contains 12 accepting states. $F = \{2, 4, 10, 11, 12, 13, 14, 15, 19, 22, 24, 25\}$.

  Each accepting state indicates that a token has been found and also represents different rules or combination of rules applicable on that token. So when we reach to an accepting state a token is generated and decide which rule or combination of rules are applicable on this token. The applicable rules for the accepting states are given in Table XI. The rules are discussed earlier in Section II (D).

TABLE XI. ACCEPTING STATES AND THEIR DESCRIPTION

Accepting States	Description
2	Single Bangla character is replaced by equivalent Braille cell
4	A consonant is replaced and then Braille cell dot 1 is inserted
10	All Bangla characters are replaced by equivalent Braille cells
11	Braille cell dot 4, 6 is inserted and three Bangla characters are replaced by equivalent Braille cells.
12	Braille cell dot 4 is inserted and then two Bangla characters are replaced by equivalent Braille cells.
13	Single character ending with hasanta are replaced
14	Braille cell dot 4 is inserted and then two Bangla characters ending with hasanta are replaced by equivalent Braille cells.
15	Braille cell dot 4, 6 is inserted and three Bangla characters ending with hasanta are replaced by equivalent Braille cells.
19	Numbers are replaced in equivalent Braille cells with number prefix inserted at first.
22	Braille cell dot 4 is inserted and then two Bangla characters are replaced by equivalent Braille cells and Braille cell 1 is inserted at last.
24	Braille cell dot 4, 6 is inserted and then three Bangla characters are replaced by equivalent Braille cells and Braille cell 1 is inserted at last.
25	Braille cell dot 4, 6 is inserted and then four Bangla characters are replaced by equivalent Braille cells.

- δ is the transition function that takes a state and an input symbol as arguments and returns a state. In a transition diagram each transition from one state to another or same state for a specific input symbol also represents a transition function. It means transition function $\delta(a,x) = b$ is a transition from state a to state b for input symbol x in transition diagram where state a and b can be same or different state. In the proposed DFA, we have 182 (7 input symbols X 26 states) transition functions. All these transition functions are represented as transition diagram in Figure 3.

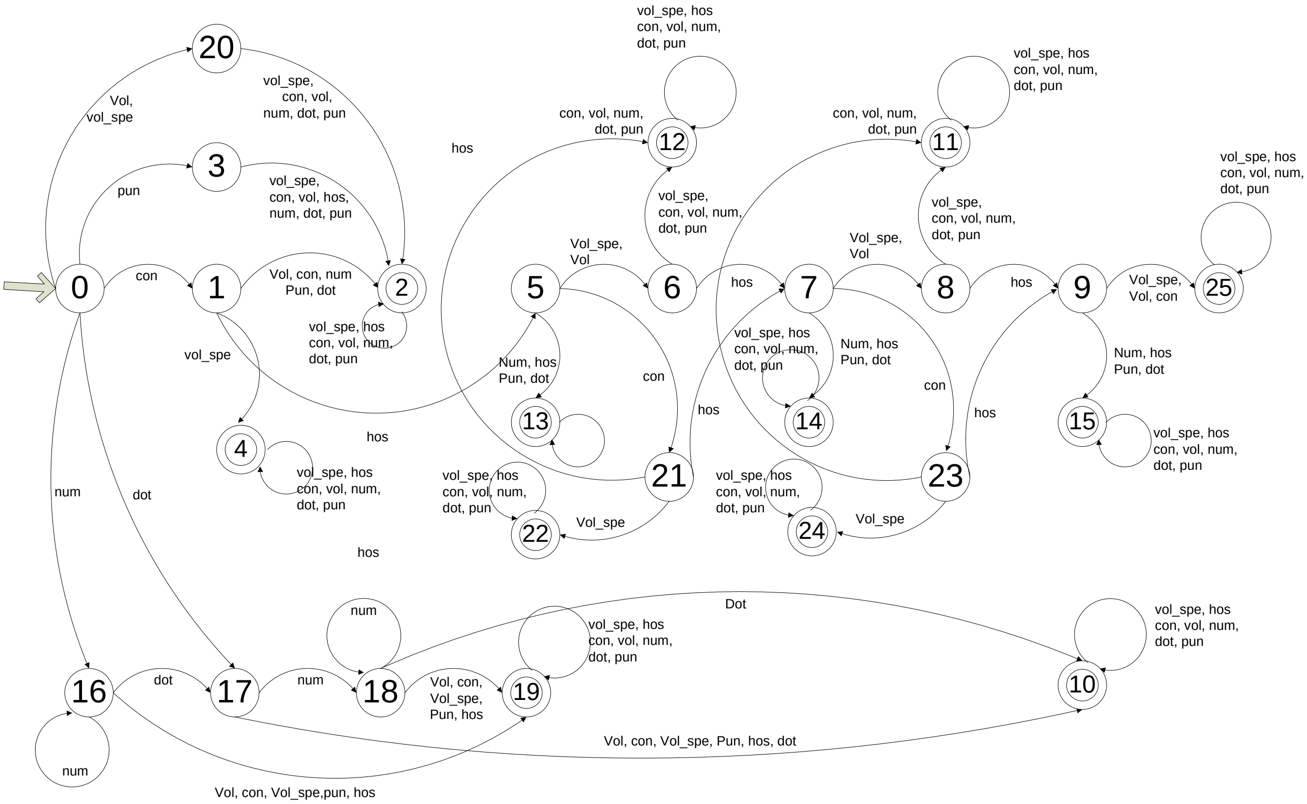


Figure 3: Transition diagram for proposed DFA

C. *Translator*

Translator maintains a look up Table of character mapping and rules. Mapping Table contains mapping between Bangla and Braille character. Rules Table contains rule number and corresponding actions. Translator receives token and rules number from Rule Engine. Then it looks up rules Table by using rule number and find out corresponding action. Finally it executes this action using Braille and Bangla character mapping Table and generates Braille symbols of Bangla text.

D. *Braille Text*

Braille text consists of equivalent Bangla Braille symbols of Bangla alphabets, digits, punctuations and formatting characters. The Braille text must be in ASCII. These Braille texts can be printed through Braille Embosser.

V. VEIW MODEL OF PROPOSED COMPUTATIONAL MODEL

View model of proposed computational model shows how the model translates Bangla text to Braille. Here an example is illustrated how a Bangla word ভ্রম is translated to its corresponding Braille representation. At first the word is passed to the rule engine. Rule engine is actually a Deterministic Finite Automata that generates a token (ভ ্ র) and an accepting state number (12) which represents a rule number. Then the token and rule number is passed to Translator. Using the rule number, translator looks up for corresponding action for the rule number. Here the action is insert dot 4 (⠈) and replace two characters (ভ,র). Then translator inserts dot 4 and looks up the Braille symbols of "ভ" and "র" in Mapping Table for replacing. Finally Braille symbols for ভ্র is generated. Similarly, using iterative way the Bangla text ভ্রম is translated into its Braille representation. This view model is shown in Figure 4.

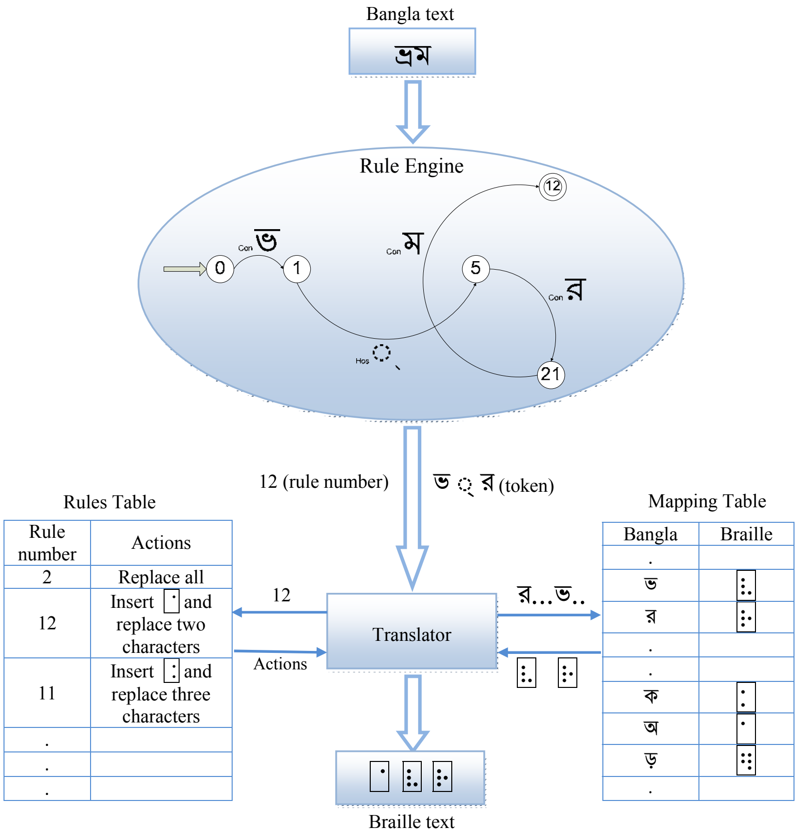


Figure 4: View Model of Proposed Computational Model

VI. CONCLUSION AND DISCUSSION

In this research work, we have analyzed and experimented with the required transformation of different rules necessary for the machine translation of Bangla to Braille. We have designed and optimized the Deterministic Finite Automata based on the experimental observations. This proposed computational model is expected to be the core rule engine for the purpose of machine translation.

As it is an absolute necessity to build an automated system through which Bangla text can be converted into Braille without the presence of any Braille expert, in the future this computational model will be implemented and tested in an environment suitable for visually impaired persons.

REFERENCES

[1] World Wide Web at http://en.wikipedia.org/wiki/Bengali_language, 22nd April 2010.

[2] The summary Report of The Bangladesh National Blindness & Low Vision Survey (30 and above age group), May 2003.

[3] World Wide Web at http://code.google.com/p/liblouis, 6th July 2010.

[4] Alasdair King, “Text and Braille computer translation” A dissertation submitted to the University of Manchester Institute of Science and Technology for the degree of Master of Science, 28 September 2001.

[5] Xuan Zhang, Ortega-Sanchez, C., Murray, l., “A System for Fast Text-to-Braille Translation Based on FPGAs” IEEE Explore, pp. 125-130, 28-26 Feb. 2007.

[6] Manzeet Singh, Prateek Vatia, “Automated Conversion of English and Hindi Text to Braille Representation” International Journal of Computer Applications 4(5):18–24, July 2010.

[7] Abdallah M Abualkishik, Dr. Khairuddin Omar, “Quran Vibrations in Braille Code” International Conference on Electrical Engineering and Informatics, pp. 12-17, 5-7 August 2009, Selangor, Malaysia.

[8] World Wide Web at http://www.webelmediatronics.in/systemVisualImpaired.htm, 15th January 2010.

[9] A Lahiri, S. J. Chattopadhyay and Anupam Basu, “A Comprehensive Indian languages tool set for the blind” 7th International Conference (ACM) Conference) on Computers and Accessibility, ISBN:1-59593-159-7, NY, USA, 2005.

[10] M. Rajasenathipathi, Dr. M. Arthanari, M. Sivakumar, “Conversion of English Text to Braille Code vibration signal for Visually Impaired People”, International Journal of Computer Science and Information Security Publication, Vol. 8 No. 5, Paper 27071032, pp. 59-63, August 2010.

[11] Satoshi Ono, Yoshinobu Hamada, Yoshitsugu Takagi, Seiichi Nishihara and Kazunori Mizuno, “Interactive Japanese-to-Braille Translation Using Case-Based Knowledge on the Web ”, PRICAI 2000 Topics in Artificial Intelligence, Volume 1886/2000, pp. 638-646, 2000.

[12] Omar Khan Durrani, K.C.Shet, “A new Architecture for Braille transcription from optically recognized Indian languages”, International Conference, Vision 2005, London. U.K.

[13] World Wide Web at http://en.wikipedia.org/wiki/Braille, 11th January 2010.

[14] প্রমিত বাংলা ব্রেইল By Blind Education and Rehabilitation Development Organisation (BERDO) and Bangladesh Blind Mission (BBM), 2001.

[15] John E. Hopcroft, Introduction to Automata Theory, Languages and Computation, second edition.

View publication stats